STATS 509 Homework 5 Xiaofeng Nie

1. Solution:

(a) Only when two random variables are linear correlated could Pearson correlation be 1, while X and Y are not linear correlated. Thus, Pearson correlation will be less than 1. On the contrary, when X and Y have monotonic relationship, Spearman correlation will be 1. In this case, when X belongs to [0,1], Y is monotonic increasing, so Spearman correlation is 1.

(b)

$$\lambda_\ell = \lim_{q\downarrow 0} P(F_1(X_1) \le q | F_2(X_2) \le q)$$

$$= \lim_{q\downarrow 0} \frac{P(F_1(X_1) \le q, F_2(X_2) \le q)}{P(F_2(X_2) \le q)}$$

$$= \lim_{q\downarrow 0} \frac{C(q,q)}{q} = \lim_{q\downarrow 0} \frac{(2q^{-\theta} - 1)^{-\frac{1}{\theta}})}{q}$$

$$= \lim_{q\downarrow 0} \frac{(2 - q^{\theta})^{-\frac{1}{\theta}} q}{q} = \lim_{q\downarrow 0} (2 - q^{\theta})^{-\frac{1}{\theta}} = 2^{-\frac{1}{\theta}}$$

$$\lambda_u = \lim_{q\uparrow 1} P(F_1(X_1) \ge q | F_2(X_2) \ge q)$$

$$= \lim_{q\uparrow 1} \frac{P(F_1(X_1) \ge q, F_2(X_2) \ge q)}{P(F_2(X_2) \ge q)} = \lim_{q\uparrow 1} \frac{(1-q) + (1-q) - (1 - C(q,q))}{1-q}$$

$$= 2 - \lim_{q\uparrow 1} \frac{1 - C(q,q)}{1-q} = 2 - \lim_{q\uparrow 1} \frac{1 - (2q^{-\theta} - 1)^{-\frac{1}{\theta}}}{1-q}$$

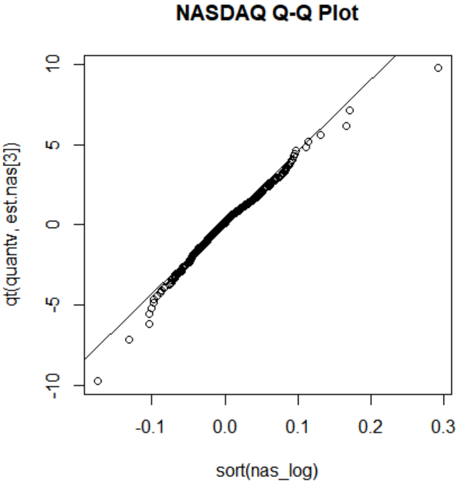
$$\textit{According to L'Hospital's rule, we have}$$

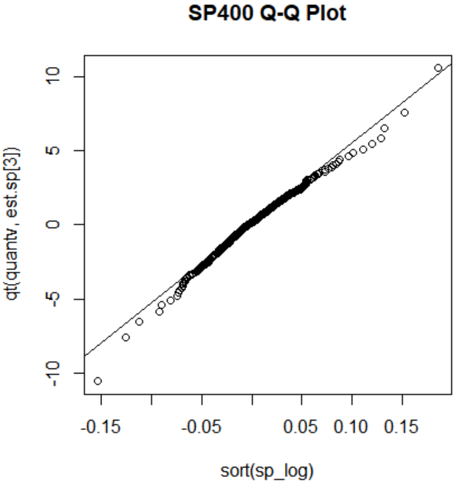
$$= 2 - \lim_{q\uparrow 1} \frac{\left(-\frac{1}{\theta}(2q^{-\theta} - 1)^{-\frac{1}{\theta}-1}\right) 2(-\theta) q^{-\theta-1}}{-1} = 0$$

2. Solution:

(a) Fitting separately, we know the mean, the standard deviation and degree of freedom for t-distribution of NASDAQ are -0.00335, 0.02302, 3.67436. he standard deviation and degree of freedom for t-distribution of SP400 are-0.00308, 0.01845, 3.47233.

```
> est.nas;est.sp
[1] -0.003350693  0.023018961  3.674359674
[1] -0.003084286  0.018450673  3.472334962
```



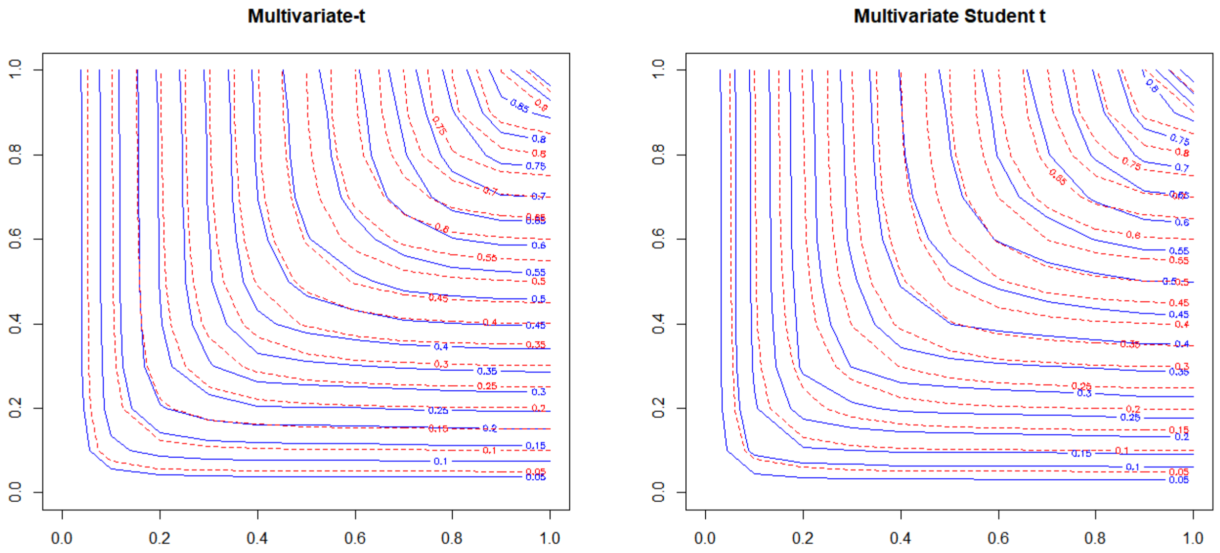


Fitting a t-copula to the data, the degree of freedom is 3.8302, and Spearman correlation is 0.8883.

Comparing with the estimated multivariate t-distribution from Problem 3-(b) from Homework 4 (df=2.8), the degree of freedom for t-copula becomes larger. Looking at the CDF plot, the left plot is the old one, which seems better than the CDF plot of t-copula (on the right). At the same time, AIC of t-copula is larger than AIC of multivariate-t distribution in HW4.

Thus, we can say that based on given datasets, t-copula based distribution is worse than multivariate-t distribution in HW4.

```
> summary(ft)
Call: fitCopula(copula, data = data, start = ..2, optim.method = "L-BFGS-B")
Fit based on "maximum pseudo-likelihood" and 1081 2-dimensional observations.
t-copula, dim. d = 2
      Estimate Std. Error
rho.1   0.8883      0.007
df      3.8302         NA
The maximized loglikelihood is 849.4
Optimization converged
Number of loglikelihood evaluations:
function gradient
      13       13
```




```
> AIC_t_copula+AIC_nas+AIC_sp
[1] -11069.22
```

```
> AIC_t
[1] -11095.61
```

3. Solution:

(a) To minimize VaR, w=-0.162, i.e. sell 16.2% NASDAQ and buy 116.2% SP400. And minimum VaR is 0.09697497 of total portfolio value.

```
> c(wmax,VaR)
[1] -0.16200000  0.09697497
```

```
> exp_shortfall
[1] 0.1415868
```

(b) To minimize variance, w=-0.11, i.e. sell 11% NASDAQ and buy 110% SP400. The minimum variance is 0.0007493792.

```
> c(w_var,var_min)
[1] -0.1100000000  0.0007493792
```

(c) The probability that the returns of the assets will simultaneously both be below their respective, relative VaRs at q = 0.003 is 0.00187.

```
> P
[1] 0.00187
```

4. Solution:

(a) We need to solve $\max_{w\in[0,1]} Sharp\ Ratio$.

$$Sharp\ Ratio = \frac{ER_p - \mu_f}{\sigma_p} = \frac{w\mu_1 + (1-w)\mu_2 - \mu_f}{\sqrt{w^2\sigma_1^2 + (1-w)^2\sigma_2^2 + 2w(1-w)\rho_{12}\sigma_1\sigma_2}}$$

In order to solve $\max_{w\in[0,1]} Sharp\ Ratio$, we need take differentiate Sharp Ratio with respect to w.

$$\frac{\partial Sharp\ Ratio}{\partial w} = (\mu_1 - \mu_2)(w^2\sigma_1^2 + (1-w)^2\sigma_2^2 + 2w(1-w)\rho_{12}\sigma_1\sigma_2)^{-\frac{1}{2}} + \left(w\mu_1 + (1-w)\mu_2 - \mu_f\right)\left(-\frac{1}{2}\right)(w^2\sigma_1^2 + (1-w)^2\sigma_2^2 + 2w(1-w)\rho_{12}\sigma_1\sigma_2)^{-\frac{3}{2}}(2\sigma_1^2 w + 2\sigma_2^2 w - 2\sigma_2^2 + 2\rho_{12}\sigma_1\sigma_2 - 2\rho_{12}\sigma_1\sigma_2 w) = 0$$

$$\Leftrightarrow (\mu_1 - \mu_2) = \frac{1}{2}\left(w\mu_1 + (1-w)\mu_2 - \mu_f\right)(w^2\sigma_1^2 + (1-w)^2\sigma_2^2 + 2w(1-w)\rho_{12}\sigma_1\sigma_2)^{-1}(2\sigma_1^2 w + 2\sigma_2^2 w - 2\sigma_2^2 + 2\rho_{12}\sigma_1\sigma_2 - 2\rho_{12}\sigma_1\sigma_2 w)$$

$$\Leftrightarrow 2(\mu_1 - \mu_2)(w^2\sigma_1^2 + (1-w)^2\sigma_2^2 + 2w(1-w)\rho_{12}\sigma_1\sigma_2) = \left(w\mu_1 + (1-w)\mu_2 - \mu_f\right)(2\sigma_1^2 w + 2\sigma_2^2 w - 2\sigma_2^2 + 2\rho_{12}\sigma_1\sigma_2 - 2\rho_{12}\sigma_1\sigma_2 w)$$

$$\Leftrightarrow 4\rho_{12}\sigma_1\sigma_2(\mu_1 - \mu_2) - 2\mu_1\sigma_2^2 w + 2\mu_1\sigma_2^2 + 2\mu_2\sigma_2^2 w + \mu_1\rho_{12}\sigma_1\sigma_2 w + \mu_2\rho_{12}\sigma_1\sigma_2 w = 2\mu_2\sigma_1^2 w + 2\mu_2\sigma_2^2 w - 2\mu_f\sigma_1^2 w - 2\mu_f\sigma_2^2 w + 2\sigma_2^2\mu_f + 2\mu_f\rho_{12}\sigma_1\sigma_2 w$$

$$\Leftrightarrow 4\rho_{12}\sigma_1\sigma_2(\mu_1 - \mu_2) + 2\mu_1\sigma_2^2 - 2\sigma_2^2\mu_f = \left(2\mu_2\sigma_1^2 + 2\mu_2\sigma_2^2 - 2\mu_f\sigma_2^2 + 2\mu_1\sigma_2^2 - \mu_1\rho_{12}\sigma_1\sigma_2 - \mu_2\rho_{12}\sigma_1\sigma_2 + 2\mu_f\rho_{12}\sigma_1\sigma_2\right)w$$

$$\Leftrightarrow w = \frac{2\rho_{12}\sigma_1\sigma_2(\mu_1 - \mu_2) - (\mu_1 - \mu_f)\sigma_2^2}{\left(\mu_2 - \mu_f\right)\sigma_1^2 + \left(\mu_1 - \mu_f\right)\sigma_2^2 - (\mu_1 + \mu_2 - 2\mu_f)\rho_{12}\sigma_1\sigma_2}$$

Thus, the weight w_T for the tangent portfolio is

$$w_T = \frac{2\rho_{12}\sigma_1\sigma_2(\mu_1 - \mu_2) - (\mu_1 - \mu_f)\sigma_2^2}{\left(\mu_2 - \mu_f\right)\sigma_1^2 + \left(\mu_1 - \mu_f\right)\sigma_2^2 - (\mu_1 + \mu_2 - 2\mu_f)\rho_{12}\sigma_1\sigma_2}$$

(b) It is true when R_p is a net return or a gross return.

For net return:

$$Net\ Return = \frac{\sum_{i=1}^N X_i^t - \sum_i^N X_i^{t-1}}{\sum_i^N X_i^{t-1}} = \frac{\sum_i^N (X_i^{t-1} R_i)}{\sum_i^N X_i^{t-1}} = \sum\nolimits_{i=1}^N \left(\frac{X_i^{t-1}}{\sum_i^N X_i^{t-1}} R_i\right) = \sum\nolimits_{i=1}^N w_i R_i$$

where X_i^t is price of i^{th} asset at time t, and R_i is the net return of i^{th} asset.

For gross return:

$$Gross\ Return = \frac{\sum_i^N X_i^t}{\sum_i^N X_i^{t-1}} = \frac{\sum_i^N (X_i^{t-1} R_i)}{\sum_i^N X_i^{t-1}} = \sum\nolimits_{i=1}^N \left(\frac{X_i^{t-1}}{\sum_i^N X_i^{t-1}} R_i\right) = \sum\nolimits_{i=1}^N w_i R_i$$

where X_i^t is price of i^{th} asset at time t, and R_i is the gross return of i^{th} asset.

Theoretically, it is not true when R_p is a log return. But their value will be similar, especially when net return is close to 0.

$$Log\ Return = \ln\left(\frac{\sum_i^N X_i^t}{\sum_i^N X_i^{t-1}}\right) = \ln\left(\sum\nolimits_{i=1}^N w_i R_i\right)$$

where X_i^t is price of i^{th} asset at time t, and R_i is the gross return of i^{th} asset.

R code for problem 2

```
library(mvtnorm)
library(copula)
library(MASS)
library(fCopulae)
library(fGarch)
library(mnormt)

nasdaq = read.csv("O:\\18WIN\\STATS 509\\HW4\\Nasdaq_wklydata_92-
12.csv",header = T)
sp400 = read.csv("O:\\18WIN\\STATS 509\\HW4\\SP400Mid_wkly_92-
12.csv",header = T)
nas_adjclose = nasdaq$Adj.Close
sp_adjclose = sp400$Adj.Close
nas_log = diff(log(nas_adjclose))
sp_log = diff(log(sp_adjclose))
len = length(nas_log)

# estimate seperate t distribution
est.nas = as.numeric(fitdistr(nas_log,"t")$estimate)
est.sp = as.numeric(fitdistr(sp_log,"t")$estimate)
est.nas;est.sp
par(mfrow=c(1,2))
quantv = (1/len)*seq(.5,len-.5,1)
qqplot(sort(nas_log),qt(quantv,est.nas[3]),main="NASDAQ Q-Q Plot")
abline(lm(qt(c(.25,.75),est.nas[3])~quantile(nas_log,c(.25,.75))))
qqplot(sort(sp_log),qt(quantv,est.sp[3]),main="SP400 Q-Q Plot")
abline(lm(qt(c(.25,.75),est.sp[3])~quantile(sp_log,c(.25,.75))))
est.nas[2] = est.nas[2]*sqrt(est.nas[3]/(est.nas[3]-2))
est.sp[2] = est.sp[2]*sqrt(est.sp[3]/(est.sp[3]-2))
# fit t-copula
data =
cbind(pstd(nas_log,mean=est.nas[1],sd=est.nas[2],nu=est.nas[3]),pstd(sp_
log,mean=est.sp[1],sd=est.sp[2],nu=est.sp[3]))
cor_tau = cor(data[,1],data[,2])
cop_t = tCopula(cor_tau,dim=2,dispstr="un",df=4)
ft = fitCopula(cop_t,optim.method="L-BFGS-
B",data=data,start=c(cor_tau,5))
summary(ft)
# empirical and theoretical cdf
dem = pempiricalCopula(data[,1],data[,2])
contour(dem$x,dem$y,dem$z,main="Multivariate Student
t",col='blue',lty=1,lwd=1,nlevel=20)
ct = tCopula(ft@estimate[1],dim=2,dispstr = "un",df=ft@estimate[2])
utdis = rCopula(100000,ct)
demt = pempiricalCopula(utdis[,1],utdis[,2])
contour(demt$x,demt$y,demt$z,col='red',lty=2,lwd=1,add=TRUE,nlevel=20)
# AIC
AIC_t_copula = -2*ft@loglik+2*2
AIC_nas = -2*fitdistr(nas_log,"t")$loglik+2*2
AIC_sp = -2*fitdistr(sp_log,"t")$loglik+2*2
AIC_t_copula+AIC_nas+AIC_sp
```

R code for problem 3

```
# Problem 3-a,b
set.seed(2015)
uvsim = rCopula(1e5,ct)
w = seq(-1,1,0.001)
n = length(w)
VaRv=rep(0,n)
var=rep(0,n)
data_sim =
cbind(qstd(uvsim[,1],mean=est.nas[1],sd=est.nas[2],nu=est.nas[3]),qstd(u
vsim[,2],mean=est.sp[1],sd=est.sp[2],nu=est.sp[3]))
for(i in 1:n){
```

```
  datat = w[i]*data_sim[,1]+(1-w[i])*data_sim[,2]
  VaRv[i] = -quantile(datat,0.005)
  var[i] = sd(nas_log)^2*w[i]^2+sd(sp_log)^2*(1-
w[i])^2+2*sd(nas_log)*sd(sp_log)*cor(nas_log,sp_log)*w[i]*(1-w[i])
}
w_var = w[which.min(var)]
var_min = var[which.min(var)]
c(w_var,var_min)
wmax = w[which.min(VaRv)]
VaR = VaRv[which.min(VaRv)]
c(wmax,VaR)
shortfall = 0
count = 0
data_sim_2 = wmax*data_sim[,1] + (1-wmax)*data_sim[,2]
n = length(data_sim_2)
for (i in 1:n){
  if (data_sim_2[i] > VaR){
    shortfall = shortfall + data_sim_2[i]
    count = count + 1
  }
}
exp_shortfall = shortfall / count
exp_shortfall

# 3-(c)
VaR_nas = qstd(0.003, est.nas[1], est.nas[2], est.nas[3])
VaR_sp = qstd(0.003, est.sp[1], est.sp[2], est.sp[3])
return_nas = data_sim[,1]
return_sp = data_sim[,2]
count = 0
for (i in 1:1e5){
  if ( return_nas[i] < VaR_nas && return_sp[i] < VaR_sp){
    count = count + 1
  }
}
P = count/1e5
P
```